DOCUMENT RESUME

ED 206 629 TM 810 343

AUTHOR Rudman, Herbert C.; And Others
TITLE Integrating Assessment with Instruction: A Report.
PUB DATE Apr 81
NOTE 75p.: Papers presented at the Annual Meeting of the National Council on Measurement in Education (Los Angeles, CA, April, 1981).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Curriculum Evaluation; *Decision Making; Elementary Secondary Education; *Information Utilization; *Inservice Teacher Education; *Material Development; Measurement; Program Implementation; Student Evaluation; Teacher Attitudes; Teacher Workshops
IDENTIFIERS *Data Box; Dependents Schools; *Test Use

ABSTRACT

The four papers presented in this symposium describe the first year and a half of a two-year study designed to examine and increase the linkage between assessment and instruction in classroom decision-making. The first year of the project was devoted to a review of the literature, development of a taxonomy interrelating measurement concepts essential to instructional decision-making, identifying contexts in which interpretation of measures of learning could be helpful, and examining tangential information related to home environments. The result was the development of The Data Box, and integrated set of materials designed to help teachers investigate the use of assessment data in instructional decision-making and ultimately to increase their own use of assessment data. Pilot data on a field review of The Data Box and data gathered at workshops indicated positive attitudes of teachers toward the Data Box and positive ratings of the workshops. A post-implementation evaluation will be carried out in May 1981. The level of teacher morale and the self-perceived needs of teachers had a great influence on the reception of The Data Box by teachers. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED206629

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Integrating Assessment With Instruction: A Report

Herbert C. Rudman, William A. Mehrens, Donna S. Wanous
and Christopher M. Clark

Michigan State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. Rudman
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 343 (1 of 4)

Part of a symposium on integrating assessment with instruction presented at the NCME annual meeting. Los Angeles, California, April 1981.

INTEGRATING ASSESSMENT WITH INSTRUCTION: A REPORT

Herbert C. Rudman, William A. Mehrens, Donna S. Wanous
and Christopher M. Clark

The purpose of this symposium will be to share the results of the first year and a half of a two-year study designed to examine and increase the linkage between assessment and instruction in classroom decision-making. We have come a long way since we reported the results of our first phase of this project, and it is our intention to share with you just how far we have come.

Last year we reported the rationale for our developmental work and shared some of the results of our review of a literature that spanned sixty years of test use. (Rudman, Kelly, Wanous, Mehrens, Clark, and Porter, 1980). As any good serial should do, we will now begin from where we last were. I will, in a sense, give a short reprise so that we can pick up the threads of this project, and my colleagues will carry on with the more exciting part of the narrative. I'm told that project leaders must be noble; I'm noble and they're having the fun.

Those of you who were with us last year may recall that we responded to an RFP that was issued jointly by the National Institute for Education and the Department of Defense Dependents Schools. These two organizations shared a common interest; the improvement of the interpretation and use of assessment information in the making of classroom decisions. The charge to those who chose to respond to the RFP was a straight-forward one; review the literature, report comprehensively on what we already know about the

state of use of tests and other assessment measures, and develop a set of instructional materials and delivery systems to disseminate the materials.

The first year of our project was devoted to searching a massive array of material that had been written and translating that into a meaningful system for helping teachers more effectively use information in making instructional decisions.

By the end of the first year we had (1) produced a monograph which reported the literature, (2) developed a taxonomy that interrelated measurement concepts that we believed were essential to instructional decision-making, (3) identified contexts in which the interpretation of formal and informal measures of learning could be helpful, and (4) examined tangential information related to home environments. We developed a system of instructional materials and implementation procedures which we called The Data Box.

At the end of that first year we had evaluated the materials with the help of teachers and school administrators and were ready to face the second year with a system that was ready to be implemented in schools.

What Have We Learned From The Past?

Earlier in these remarks, I alluded to a review of the literature which dealt with the interpretation and use of assessment information. This review covered 229 studies and other publications during the years, 1922-1980. The review was organized around four questions; (1) how do teachers use test information for instructional decisions? (2) what has research indicated about the linkage between testing and teaching? (3) what is known about the successful practice of in-service training of

teachers? and (4) what is and has been available to teachers to aid them in interpreting and using assessment information for instructional decision-making?

Our review revealed a literature that is uneven in quality, in scope of serious study, filled with inconsistencies, and difficult to find in the traditional channels of research. Some examples of the answers or near-answers to the four organizing questions were the following:

1. Teacher's estimates of pupil performance on tests are remarkably stable and consistent. They are not nearly as prone to expectation set as earlier characterized (Morine-Dershimer, 1979; Sorotzkin, Fleming and Anttonen, 1974; Beggs, Mayer and Lewis, 1972; Pedulla, Airasian, Madaus, and Kellaghan, 1977; Rosenthal and Jacobson, 1968).

2. It would appear that there is a gap between what those most closely affected by testing think and what others write about what they think (Stetz and Beck, 1977, 1978, 1979; Goslin, 1967; Rudman, 1977).

3. In contrast to teachers' perceptions of their students' test scores, there is some evidence that teachers' reporting of their students' classroom interpersonal behavior is neither stable nor accurate (Elmore and Beggs, 1972; Barnhard, Zimbardo and Sarason, 1968; Openshaw, 1967; Feshbach, 1969; Tolor, Scarpetti, and Lane, 1967).

4. There is conflicting evidence concerning teachers' sensitivity to the information they use in instruction. Some researchers report that teachers are sensitive to the reliability of the information they receive and can adjust their judgments of students accordingly; teachers differentiate among the data they receive and use them differentially depending upon the instructional issue they confront (Shavelson, Cadwell and Izu, 1977).

5. There is considerable reason to believe that teachers have limited knowledge about fundamental measurement concepts (Goslin, 1967; Farr and Griffin, 1973; Noll, 1955, 1961; Olejnik, 1979).

6. Recent writing would indicate that the use of criterion-referenced tests are preferred over that of norm-referenced tests, but with the exception of one article reviewed, most of the literature calling for the use of criterion-referenced tests is prescriptive rather than descriptive (Yeh, 1978; Popham and Husek, 1969; NEA 1975; Airasian and Madaus, 1972).

7. Those studies that were descriptive tended to show a preference for norm-referenced tests and the standard scores in which the results are couched. (Yeh, 1978; Olejnik, 1979; O'Reagan et. al., 1979; Nearine, 1970).

8. In contrast, only 35 of teachers quizzed, favored increased use of criterion-referenced tests. (Beck and Stetz, in press).

9. Most in-service experiences are similar in design; lectures, self-study modules and pamplets, analysis of video-tape incidents, and are usually structured for group participation and study (NEA, 1976, 1977a, 1977b, 1977c, 1977d; Deanda, 1975; Capie, 1978; Harrison, 1976).

One last observation; after prolonged discussion of the information synthesized from the review of the literature, the staff agreed with the following eight principles which served to guide our development of The Data Box and the subsequent implementation of it. (Cruikshank, Lorish and Thompson, 1979). An in-service program can be most effective if it is:

1. predicated on a longer term plan as opposed to a single-session meeting.

2. one that utilizes personal contact between the trainer and those trained.

3. a program that utilizes on-site contact between teachers and trainer, rather than a program conducted away from the site where a reasonable opportunity exists for application of skills and concepts learned.

4. based upon collaboration between teachers, administrators and those specialists charged with conducting the in-service program.

5. structured so that there is an opportunity for the real world of the classroom to intermesh with those materials, concepts and skills to be learned.

6. predicated on a follow-up program which utilizes local staff and specialists who may be external to the school and/or system.

7. based upon a plan that utilizes demonstration, supervised trials and feedback.

8. one that utilizes the opportunity for teachers to relate to other teachers and who may serve as consultants to their colleagues.

What Have We Learned to Date?

While much of what will follow will detail the contents of The Data Box, what we learned from our experiences with the administrator and teacher workshops, and the evaluation of the instructional materials and training sessions, I should like to end this presentation with some tentative "lessons" we have drawn that we believe have some implications for the use of data in classrooms.

1. Some teachers are uncomfortable dealing with ambiguities; they prefer to be given a solution they can accept or reject rather than be asked to deduce what a reasonable solution to a problem may be. The Data Box has been developed with the assumption that teachers have solutions to problems that may be every bit as good as those offered by The Data Box. For those teachers who are comfortable with alternative approaches to the solution of a problem, The Data Box is welcomed as a useful device. For those who have little patience with teasing out a solution, The Data Box is viewed less positively, but usually with the caveat that it has great promise for the young person still in teacher preparation programs

2. The Data Box lends itself best to prolonged inservice programs rather than one-day orientation sessions. There is reason to believe that the most that a one-day experience with The Data Box can bring is a heightened sensitivity to the need for further training in the use of assessment information in the instructional process.

3. If an inservice program is to be planned for teachers it must include teachers on the planning committee and its objectives must be clearly enunciated well in advance of the inservice program. Differences in expectations of those who attend the program and those who offer it can be counter-productive.

4. While there is reason to believe that The Data Box is a useful tool to upgrade teachers' interpretation skills and their knowledge of measurement concepts, it clearly is no panacea. While the results comparing The Data Box with and without inservice preparation have not yet been analyzed chances are good that the most effective use of the materials can be obtained in connection with some informed leadership in the initial stages of their use. The MSU team tend to believe that informed leadership and directed discussion and instruction prior to small group use of The Data Box is better than the "friendly exchange of uninformed opinion."

5. The experience of the MSU team in Germany led us to believe that more teachers have had measurement training than we might have anticipated. Nevertheless we believe that teachers can use some "refreshers" on the most relevant of the measurement concepts from their ability to understand and interpret the information and format of documents they use.

It could be that some teachers may assume that because they use documents with some regularity that they are interpreting them accurately.

6. The administrative/leadership styles differed considerably from one building in Germany to the next. An interesting studywould be the relationship between teachers' interest and competency in using information in their instructional decision-making, and the leadership styles of the building administrator with whom they work.

7. Teacher connected with this project showed considerable interest in the use of standardized tests; there was a concern among teachers in the Department of Defense Dependents Schools for more rather than less standardized achievement testing.

8. The MSU team was impressed with the need expressed by administrators and teachers for a clarification of the awards to be given for taking part in professional development activities. We were impressed with the similarity of the schools we were in with those we encountered in the United States; a similarity that led us to believe that a clear identification of a professional awards system was an integral part of motivating teachers to tak a role in their own professional development of their assessment skills.

REFERENCES

Airasian, Peter W., and George F. Madaus. "Criterion-Referenced Testing in the Classroom." Measurement in Education 3 (1972): 1-8.

Barnard, J.W., P.G. Zimbardo, and S.B. Saranson. "Teachers' Ratings of Student Personality Traits as They Relate to IQ and Social Desirability." Journal of Educational Psychology 59 (1968): 128-132.

Beggs, Donald G., Roy Mayer, Ernest L. Lewis. "The Effects of Various Techniques of Interpreting Test Results on Teacher Perception and Pupil Achievement." Measurement and Evaluation in Guidance 5 (April 1972): 290-297.

Capie, William et. al. Teacher Performance Assessment Instruments: Plans for Practice Rating. Atlanta: Georgia State University, 1978. (Ed 146539)

Cruikshank, D., C. Lorish, and L. Thompson. "What We Think We Know About Inservice Education." Journal of Teacher Education 30 (1979): 27-32.

Deanda, Natividad. Helping Children Develop Healthy Self-Concepts. Washington, D.C.: National Institute of Education, 1975. (Ed 129455).

Elmore, Patricia and Donald L. Beggs. "Stability of Teacher Ratings of Pupil Behavior in a Classroom Setting." Paper presented at a meeting of the American Personnel and Guidance Association, March 1972). (Ed 0666855).

Farr, Roger and Margaret Griffin. "Measurement Gaps in Teacher Education." Journal of Research and Development in Teacher Education 7 (1973): 19-28.

Feshbach, N.D. "Student Teacher Preferences for Elementary School Pupils Varying in Personality Characteristics." Journal of Educational Psychology 60 (1969): 126-132.

Goslin, D.A. Teachers and Testing. New York: Russell Sage Foundation, 1967.

Harrison, Susan et. al. Information for Inservice Training Development: A Collection from Field Efforts of the Rocky Mountain Regional Resource Center. Washington, D.C.: Bureau of Education for the Handicapped, 1976. (Ed 238025).

Morine-Dershimer, Greta. "How Teachers See Their Pupils." Educational Research Quarterly 3 (Winter 1978-79): 43-53.

National Education Association. Guidelines and Cautions for Considering Criterion-Referenced Tests. Washington, D.C.: The Association, 1975.

National Education Association. Critical Moments in Teaching. Washington, D.C.: National Institute of Education, 1976. (Ed 171686).

National Education Association. Designing Effective Instruction - A New Adventure in Teaching. Washington, D.C.: National Institute of Education, 1977a. (Ed 166999)

National Education Association. Evaluation of Team Teaching and Children's Continuous Progress. Washington, D.C.: National Institute of Education, 1977b. (Ed 166158).

National Education Association. The Effective Use of Observation as a Diagnostic Technique. Washington, D.C.: National Institute of Education, 1977c. (Ed 164529)

National Education Association. Administration of a Standardized Test. Washington, D.C.: National Institute of Education, 1977d. (Ed 164530)

Nearine, Robert J. "The Test, The Time, and the Teacher." Measurement and Evaluation in Guidance 2 (Winter 1970): 214-216.

Noll, Victor H. "Requirements in Educational Measurement for Prospective Teachers." School and Society 82 (September 17, 1955): 88-90.

NSBA Research Report 1977-1. Standardized Achievement Testing. Washington, D.C.: National School Board Association, 1977, p. 53.

Olejnik, Stephen F. "Standardized Achievement Programs Viewed from the Perspective of Non-Measurement Specialists." Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, April, 1979.

Openshaw, K. "A Failure of the Minnesota Teacher Attitude Inventory to Relate to Teacher Behavior." Journal of Teacher Education 18 (1967): 233-239.

O'Regan, Myra, Peter Airasian, George Madaus. "The Use of Standardized Test Information by Irish Teachers." Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, April, 1979.

Pedulla, Joseph K., Peter Airasian, George Madaus and Thomas Kellaghan. "Proportion and Direction of Teacher Rating Changes of Pupil's Progress Attributable to Standardized Test Information." Journal of Educational Psychology 69 (1977): 702-709.

Popham, W. James and T.R. Husek. "Implications of Criterion-Referenced Measurement." Journal of Educational Measurement 6 (Spring, 1964): 1-9.

Rosenthal, Robert and Lenore Jacobson. Pygmalion in the Classroom: Teacher Expectations and Pupil's Intellectual Development. New York: Holt, Rinehart and Winston, 1968.

Rudman, Herbert C. "The Standardized Test Flap." Phi Delta Kappan. 59. (November, 1977): 179-185.

Rudman, Herbert C., Janet L. Kelly, Donna S. Wanous, William A. Mehrens, Christopher M. Clark, and Andrew C. Porter. Integrating Assessment With Instruction: A Review (1922-1980). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, Research Series No. 75, 1980.

Shavelson, Richard., Joel Cadwell, and Tonia Izu. "Teachers' Sensitivity to the Reliability of Information in Making Pedagogical Decisions." American Educational Research Journal. 14 (Spring 1977): 83-97.

Sorotzkin, Feige, Elyse Fleming, and Ralph Anttonen. "Teacher Knowledge of Standardized Test Information and Its Effects on Pupil IQ and Achievement." Journal of Experimental Education 43 (Fall 1974): 79-85.

Stetz, Frank P. Report of a Market Survey to Sample Opinions of Sales Representatives and Users Concerning The 1973 Stanford Achievement Test. Market Research Report No. 10. New York: The Psychological Corporation, 1977, p. 25.

Stetz, Frank P. and Michael D. Beck. "A Summary of Opinions Concerning Users of Educational Tests." Paper presented to the National Council on Measurement in Education, Toronto, 1978.

Stetz, Frank P. and Michael D. Beck. "Comments from the Classroom: Teachers' and Students Opinions of Achievement Tests." Paper presented to the National Council on Measurement in Education, San Francisco, 1979.

Tolor, A., W.L. Scarpetti, and P.A. Lane. "Teachers' Attitudes Toward Children's Behavior Revisited." Journal of Educational Psychology. 58 (1967): 175-180.

Yeh, Jennie P. "Test Use in Schools." Washington, D.C.: U.S. Department of Health, Education and Welfare, National Institute of Education, 1978.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Rudman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 343 (2 of 4)

The Data Box: A Description

Donna Wanous
Herbert Rudman
William Mehrens
Christopher Clark

Michigan State University

Paper presented as part of a symposium on "Integrating Assessment with Instruction: A Report" at the annual meeting of the National Council on Measurement in Education in Los Angeles, California, April, 1981.

The Data Box: A Description

Background

The Data Box is an integrated set of materials designed to have teachers investigate the use of assessment data in instructional decision-making and ultimately to increase their own use of assessment data. The Data Box consists of six components. These components are centered around a series of vignettes based on everyday classroom events in which an instructional decision must be made. Each of these components and their accompanying characteristics and assumptions will be discussed briefly in this paper.

This integrated set of materials was developed so that it would contain certain characteristics. These characteristics were considered important to the successful implementation of the Data Box. Among these characteristics were the following:

1. The Data Box needed to be a self-contained unit. A unit which could be disseminated with or without any accompanying on site professional development.

2. The vignettes within the Data Box needed to cover a wide range of classroom events in order to appeal to a cross section of teachers with different levels of skills and knowledge.

3. Within the Data Box there needed to be a provision for the addition of site specific instructional situations.

In addition to these characteristics, the original development of the Data Box was based on a set of guiding assumptions. These assumptions were generated from three primary sources: 1) the findings from the review of the literature; 2) the suggestions of the External Advisory Panel; and 3) the knowledge and experiences of the staff. These assumptions will be discussed in detail in the next section entitled "Taxonomy", the first component of the Data Box.

The Components

Taxonomy. The taxonomy is the organizational guide for the Data Box (see Figure 1). In addition, it reflects the assumptions that have directed the development of these materials. These assumptions are highlighted in its two major dimensions: 1) critical classroom events, and 2) major measurement concepts.

The staff thought that it was especially important that the vignettes be based on realistic everyday classroom events that might be encountered. One way of organizing these events was to consider the various important major instructional decisions that occur during the school year. Examples of these decisions include pacing, evaluating and classifying. These decisions are referred to as "contexts" in the taxonomy.

The second dimension, "concepts," represents the major categories of measurement concepts that the staff believed were necessary to make the instructional decisions embedded in the "contexts." These "concepts" are organized into nine major categories. Within each of the nine major categories, is a more specific set of objectives.

The cells within the taxonomy contain the vignettes (by number) that represent the intersection of two dimensions. For example, if a participant wanted to know more about "relating educational decisions to data" at the "beginning of the year," he/she would select either vignette 1.1 or 1.2. Thus the taxonomy also serves as a sort of table of contents.

Vignettes. Each of the vignettes is presented on a 5" x 8" card and organized according to the "context" dimension of the taxonomy (see Figure 2). There are currently 13 vignettes. In each of the 13 vignettes a classroom situation which requires an instructional decision is described. The reader is then asked to specify what he/she would do in that situation. (A more detailed explanation of this process is provided at the end of this paper.)

CONTEXTS	Relating Educational Decisions to Data	Knowing What, Where & How to Use Data in Dec. Making	Knowing the Sources of Data	Knowing Data Collection Processes	Understanding Data Characteristics	Interaction of Data With Sources and Processes	Specifics of Data Use	Knowing How to Disseminate Information	Disseminating Information to Various Groups							
	CONCEPTS/VIGNETTES															
1.0 BEGINNING OF YEAR	1.1 1.2	1.2	1.1 1.2	1.2	1.1	1.2	1.1 1.2									
2.0 CLASSIFYING	2.1	2.1	2.1	2.1		2.1	2.1									
3.0 DIAGNOSIS	3.1 3.2	3.1 3.2	3.2	3.1	3.2	3.1 3.2	3.1									
4.0 PACING	4.1	4.1	4.1	4.1	4.1	4.1	4.1									
5.0 RETENTION AND PROMOTION	5.1	5.1	5.1	5.1	5.1	5.1	5.1									
6.0 EVALUATING	6.1 6.2	6.1 6.2	6.1	6.1	6.1 6.2	6.2	6.1 6.2	6.1								
7.0 PARENT CONFERENCE	7.1 7.2	7.1 7.2	7.1 7.2	7.1 7.2	7.1	7.1 7.2	7.1 7.2	7.1 7.2	7.1 7.2							
8.0 CURRICULUM PLANNING	8.1 8.2	8.1 8.2	8.1 8.2	8.2	8.1 8.2	8.1 8.2	8.1 8.2	8.1	8.1 8.2							

Figure 1. A Contextual Events/Measurement Taxonomy

Figure 2

Sample Vignette

2.1 **Classifying/Grouping for Mathematics Instruction**

The first marking period is just completed. You have turned in your report cards and are beginning to plan for subsequent instruction. During the second nine weeks you will have a student teacher. You have met with this person and, since she is very interested in the teaching of arithmetic, you have jointly decided to do some intraclass grouping in mathematics.

Given the 4th Grade Class List and all the information provided in *The Data Box* determine (1) how many groups you should set up, (2) who should be in which groups, and (3) what the instruction should be in those groups. Present a rationale for your answers.

In doing the above tasks what information in *The Data Box* did you use? What information not in *The Data Box* would have been helpful in determining your answers to the above questions?

In responding to this vignette assume that the 8 quizzes covered the following topics (see *The Data Box* for descriptions).

Quiz		Topic
1	Unit 1:	Numeration Parts A and B
2	Unit 1:	Numeration Parts C and D
3	Unit 2:	Addition & Subtraction Parts A, B and C of Addition
4	Unit 2:	Addition & Subtraction Part D of Addition, Parts A & B of Subtraction
5	Unit 2:	Addition & Subtraction Parts C, D, and E of Subtraction
6	Unit 3:	Measurement Parts A, B, and C
7	Unit 3:	Measurement Parts D, E, and F
8	Unit 3:	Measurement Parts G and H

After you have completed the tasks specified above, turn to page 11 in *The Interpretive Manual* and compare your responses to the sample responses provided there.

As was mentioned previously, an attempt was made to make these vignettes as realistic as possible. At present these vignettes center around a hypothetical fourth grade classroom. However the developers have found the vignettes to be general enough that they are applicable across several grade levels.

Documents. In the process of arriving at a solution to the vignettes, the reader is asked to consider a set of documents. These documents consist of both formal and informal assessment data that might be available in the typical school (see Figure 3). For example, these data include standardized achievement test results, grade book pages, reading diagnostic test results, court orders, and various memoranda regarding parental conversations.

They are divided into three categories: school characteristics, home and family information, and individual characteristics. The data are summarized at a variety of levels including classroom, building and district.

As the reader considers the various documents, he/she is asked to record the code of the documents that they have used in the decision making process. These codes are recorded on the memo pad provided in the Data Box.

Memo pad. Each Data Box contains a memo pad (see Figure 4) on which the readers are asked to record the date, the vignette with which they are working, the documents they selected and their solution. The readers are asked to record these data for each vignette they complete.

The memo pad was included in the Data Box for the following reasons:

1. To provide the reader with a convenient way to record the information he/she used throughout the decision-making process--especially when he/she compares his/her solution to that given in the interpretive manual.

2. To furnish the staff with a record of the data used in the decision-making process and the solution itself across the various readers.

Interpretive manual. After the reader has arrived at a solution, he/she is referred to the interpretive manual. The interpretive manual contains the "solutions" to the vignettes. These solutions were developed by the authors

CLASS SUMMARY

Date of Testing: 10/79
Grade : 4

STANDARDIZED ACHIEVEMENT TEST
Teacher: Mr. George Young
School : South Elementary
System : Great River

NATIONAL STANINE PROFILE 1 2 3 Below Average / 4 5 6 Average / 7 8 9 Above Average						NO TESTED
STANINE SUMMARY						
N	%	N	%	N	%	
10	37	9	33	8	30	30
6	22	11	41	10	37	30
7	26	10	37	10	37	30
6	22	10	37	11	41	30
9	33	10	37	8	30	30
8	30	9	33	10	37	30
10	37	6	22	11	41	30
9	33	10	37	8	30	30
2	8	16	59	9	33	30
8	30	11	41	8	30	30
5	19	13	48	9	33	30
7	26	8	30	12	44	30
7	26	11	41	9	33	30
6	22	12	44	9	33	30
5	19	13	48	9	33	30
7	26	11	41	9	33	30

1 2 3 Below Average	4 5 6 Average	7 8 9 Above Average	NO TESTED
6 22	16 59	5 19	30

TESTS	Number Right	Number Possible	Scaled Score	%ile Rank Local	%ile Rank Nat'l	Grade Equiv
	SCORE DETAIL					
	MEAN SCORES					
Vocabulary	25		151			4.6
Reading - Part A	41		163			5.9
Reading - Part B	40		157			5.8
Reading Comprehension	80		164			6.2
Word Study Skills	52		179			7.9
Math Concepts	25		169			7.1
Math Computation	27		163			6.0
Math Applications	20		151			5.4
Spelling	36		157			5.8
Language						
Social Science	19		163			6.3
Science	21		166			6.5
Listening Comprehension	37		148			4.8
AREA TOTALS						
Total Reading	133		167			6.0
Total Mathematics	73		166			6.5
Total Auditory	62		148			4.8
BATTERY TOTAL						
COMPLETE BATTERY	343		158			5.3

OTIS-LENNON MENTAL ABILITY TEST	SCORE DETAIL							
	Mental Age Yrs	Mental Age Mos	Raw Score	Grade Norms PR	Grade Norms S	Age Norms %ile Rank Local	Age Norms %ile Rank Nat'l	Dev I.Q.
AGE NORM STANINE ---			47					100

STANDARDIZED ACHIEVEMENT TEST

Test Level : INT. I
Form : Z

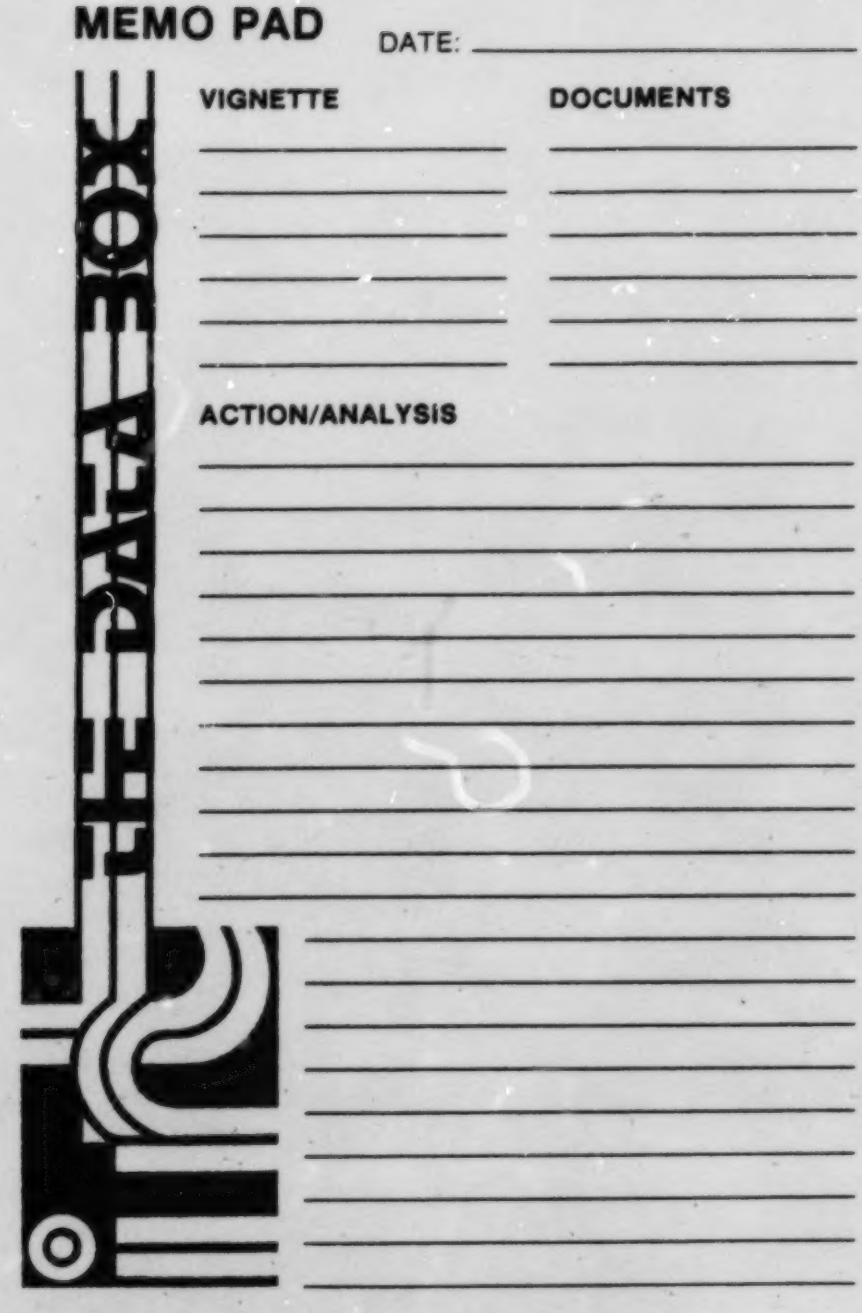
Norm Used
Grade 4.2

STANDARDIZED SCOLASTIC APTITUDE TEST

Level : Element 1 Grade 3 Beg.
Form : V

Figure 4

Sample Page of the Memo Pad



of the Data Box based on their background and knowledge. The solutions are not intended to be "the one right answer" but rather a realistic suggested solution.

Each solution contains a description of the documents used and the reasons for how the author arrived at the solution he/she did. In addition, descriptions of documents not found in the Data Box but ones that the author judged as potentially useful are identified.

Applications manual. The sixth component of the Data Box is the applications manual. There are five major parts of the applications manual.

In the first section, an overview of the purpose and components of the Data Box is provided. A brief history of the development of the Data Box and a condensed version of the most important findings from the review of the literature follow in the second part.

Section three contains a discussion of how the Data Box can be used to serve a variety of purposes. These purposes range from use by an individual teacher to a plan for a district-wide inservice program.

In section four the definitions of the relevant measurement concepts and skills are offered. These definitions were selected for inclusion by the authors because they were crucial to the solutions of the vignettes.

The total bibliography used in the review of the literature constitutes the last section. This bibliography was provided in case a reader wanted more information on a particular topic that was discussed in section two.

The Data Box Process

In an attempt to relate all the components of the Data Box to one another, a review of the process used by participants in working with the Data Box will be given. The following are the suggested steps of the process:

1. The reader chooses a vignette from those outlined in the taxonomy.
2. After reading the vignette, the reader goes to the document and selects the ones he/she judges to be relevant.

3. The reader consults the interpretive manual for the definitions of any measurement terms he/she does not understand.

4. The reader arrives at a solution. He/she then records the solution and the relevant documents on the memo pad.

5. Next the reader compares his/her solution to the solution designated in the applications manual.

6. The reader records any suggestions for revisions that the authors should make.

Summary

Since the Data Box is still considered to be in the developmental stage, it is anticipated that it will be revised based on reader reactions. To date that feedback has been positive and the recommendations have been most helpful.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Evaluating the "Integration of Assessment With Instruction" Project

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Rudman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

William A. Mehrens, Herbert C. Rudman
Donna S. Wanous and Christopher M. Clark

Michigan State University

TM 810 343 (3 of 4)

Part of a symposium on integrating assessment with instruction presented at the NCME annual meeting, Los Angeles, California, April 1981.

Evaluating the "Integration of Assessment With Instruction" Project

William A. Mehrens
Michigan State University

Previous presentors have given you some background on the project, explained the development of the materials, and discussed the workshop presentations. As you are aware, the project is not complete and final evaluation must await the May trip to Germany and the gathering of follow-up data.

Today I will present evaluative data of two types: (1) pilot data on a field review of The Data Box and (2) data gathered during and immediately following the workshops. I will refer to this second set of data as Pre-implementation data since The Data Box had really not been implemented at this time. Finally, I will describe briefly our plans for the final evaluation.

If you attended our symposium in Boston last year you may recall hearing Andy Porter speak of our anticipated evaluation activities. Subsequently we sent an evaluation plan to NIE. Those were, in my opinion, very well thought out but a bit idealistic. Real life evaluations seldom completely match the theoretical plans. At places in this presentation I will tell you our original plans, describe what we did, and explain the difference in procedure.

The original goals for the evaluation activities can be summarized in the following four questions (Porter, 1980):

1. What are the reactions of teachers and other school personnel to the format and content of the materials and two delivery systems?

2. After having experienced training through the use of materials with a delivery system, what is the state of teacher knowledge about measurement concepts and their application in a classroom setting?
3. As a result of training, what changes are evidenced in teacher classroom practices?
4. Are the trained teachers better teachers than they were prior to training?

1. Pilot Data on a Field Review of The Data Box

A. Advisory Committee Review

We have an external advisory committee which has been advising the project. In the developmental phase of The Data Box this committee was asked to review an incomplete draft set of the materials. The composition of that committee included elementary and secondary school teachers and principals, school counselors, an educational researcher, and an evaluation design specialist. We wished to learn if they

1. were in support of the measurement concepts which we had selected;
2. saw the critical incidents as ones which they deem important to teachers;
3. considered the vignettes to be clearly presented and sufficiently complete;
4. viewed the instructions for what a teacher is expected to do as understandable;
5. judged the delivery systems to be usable during the school year as a part of on-going teacher inservice programs;

6. considered the nature of the task within the boundaries of what is likely to be acceptable to teachers;
7. would expect support from other school personnel for teachers to participate in either or both of our two training programs;
8. thought our materials likely to be used by administrators and counselors as well as teachers;
9. judged the back-up information available to a teacher for use in responding to a vignette/problem typical of that which might be available to a teacher or whether additional information should have been included.
10. endorsed the descriptions of what might have been done as a "school solution" to the vignette problem as realistic and that there was adequate discussion of the measurement concepts.

In general we received positive reactions from the advisory committee. They did, of course, make several specific suggestions with respect to the draft materials. We incorporated these suggestions into our material where appropriate.

B. The Field Review

The original proposal called for a field testing of the draft materials in order to get a preliminary indication of teachers' acceptance of The Data Box as well as their perceptions of the project's utility and practicality. The field testing involved providing a sample of teachers with all the materials and collecting data on their reactions to the various components.

1. Questionnaire Construction

The actual questionnaire was constructed by Dr. Rudman with the assistance of Timothy J. Van Susteren. Details and the rationale for the

final instrument are detailed in a paper by Van Susteren (1980) and will not be repeated here. The final directions given to the teachers in the field study and the questionnaire can be found in Appendix A and Appendix B.

2. Sample

The sample was composed of 34 teachers and 3 administrators enrolled in the College of Education at Michigan State University. It was a volunteer sample with some selection made within the volunteer group to obtain some balance on several variables. A breakdown of the sample on various characteristics is shown on Table 1.

- - - - - - - - - - - - - - - -

Table 1 Here

- - - - - - - - - - - - - - - -

3. Analysis and Results

The total set of analyses and reports are presented in Van Susteren (1980). The basic results are shown here in Table 2. The results shown are the percentage of individuals who responded "strongly agree"or "agree" as well as the mean rating for each question.

- - - - - - - - - - - - - - - -

Table 2 Here

- - - - - - - - - - - - - - - -

The attitude of the sample of teachers was, in general, very favorable toward the various components of The Data Box. For example, on 42 of the 53 items in the questionnaire, the percentage of favorable responses was between 80% and 100%. There were a few notable exceptions to the pattern of general agreement. Only 43% of the sample agreed that the scores found in the document file are familiar to most teachers. However, such a response pattern serves to support the claim of the research staff that many teachers are inadequately trained in measurement (Rudman et. al. 1980). Another notable

exception was the low percentage of agreement that the information in the document file is easy to find. The response to this item indicates a need for better organization and labeling of the various documents in that file.

Coefficient alphas were calculated for gruops of questions as well as for the total set. The results are shown in Table 3. As can be seen, the total instrument is quite homogenous, the items do in general appear to be measuring over all instrument quality. Scales 2 and 3 are not very homogenous but one must keep in mind that these scales are composed of only eight and four items respectively. Further details regarding item-scale correlations can be found in Van Susteren (1980).

- - - - - - - - - - -

Table 3

- - - - - - - - - - -

Two "research" questions were investigated in the study: (1) Do teachers from rural schools react differently toward The Data Box than those from suburban or urban schools? (2) Do more experienced teachers react differently toward The Data Box than those with less experience? Although there were, as expected, differences at the .05 level on a few of the items the conclusion for both research questions was "No."

In summary, the formative evaluations of both the advisory committee review and the field review provided us with encouragement and specific suggestions for revising our materials.

II. Pre-Implementation Data (gathered during the workshops)

The evaluation data gathered during the workshops was basically of four types - gathered on three forms: a 20 item preassessment knowledge test (see Appendix C), a one page workshop evaluation form which assessed both the workshop and The Data Box materials (see Appendix D), and a set

of interview questions (see Appendix E). Each part will be discussed separately. In addition we obtained some specific feedback from the workshop participants regarding a few of the vignettes.

A. Pre-Assessment Knowledge Test

The pre-assessment test was composed of 20 questions which were related to five of the nine measurement concepts we believe teachers need to know. (The nine measurement concepts are shown in Appendix F. The compilation and rationale for this list was presented by me in Boston last year [Mehrens, 1980].) As can be seen from the Pre-Assessment Test in Appendix C, not every concept was tapped by the test, nor were the concepts tapped weighted equally. The test has five questions on Concept 4, two on Concept 5, five on Concept 6, seven on Concept 7, and one on Concept 8.

We still believe teachers need to know the nine major measurement concepts, but several are not as amenable to measuring via paper/pencil tests as by other means.

The original plan was to give the pre-assessment to the teachers/administrators at the beginning of each workshop. This was done at Karlsruhe. However, there was some feeling that this approach served somewhat to dampen the original enthusiasm for the workshop and to result in fewer volunteers to analyze The Data Box in the subsequent weeks. Therefore, in the other settings the pre-assessment was given subsequent to the workshops and only to the subset of workshop participants who volunteered for future participation. This makes any comparison of data across workshops fatuous. The pre-assessment results are shown in Table 4.

- - - - - - - - - - - - - -

Table 4 Here

- - - - - - - - - - - - - -

As can be seen, the mean's ranged from 10.17 to 12 with an overall mean of 11.05 (Max = 20, Chance = 4.6). Thus, there is considerable room for growth should The Data Box experience prove to be a useful tool for learning measurement concepts. It is interesting to note that some teachers felt they understood measurement concepts and did not need any inservice training on such material. The test scores did not corroborate such self assessment.

We should point out that we do not believe the subsequent evidence regarding cognitive gain (or lack of gain) in measurement concepts is preeminent. The experiences themselves and the awareness of using data in decision making are our primary goals for the volunteers.

B. Workshop Evaluati[illegible]rm

A one page workshop evaluation form (see Appendix D) was administered to the participants at the end of the workshops. There were two workshops for the administrators: one in North Germany and one in South Germany. The one in South Germany was scheduled for only 2 hours (rather than the full day we had planned). Since the workshop for the South Germany administrators was more of an overview of the project than a workshop we did not administer the workshop evaluation form to them. Thus, we administered the form to the North Germany administrators and the four teacher workshop attendees (mostly teachers but some school building administrators) at Rhein-Main Elementary, Hainerberg, Karlsruhe, and Heidelberg. A summary of the results is shown in Table 5.

- - - - - - - - - - - - -

Table 5 Here

- - - - - - - - - - - - -

The first six variables pertained to whether the respondents felt the workshop objectives had been met. As can be seen, the ratings were reasonably high for these six variables. (Means for total teachers ranged

from 3.1 to 4.0 on a five point scale). Measurement concepts received the lowest rating and we believe this to be accurate - we really did not dwell on measurement concepts as much as we did in attempting to instill in the participants a positive attitude toward integrating assessment with instruction.

The second six variables had to do with The Data Box materials. Ratings were lower for these variables, hovering around the neutral point for teachers but closer to 4.0 for the administrators. We think there are two primary reasons for this. (1) The teachers had an incorrect set of expectations for The Data Box. Somehow, they thought The Data Box would provide answers for specific classroom problems. It was hard for teachers to change from this initial set. (2) The workshop time was simply not sufficient to allow the participants to go over all the materials in The Data Box. Thus, they had trouble rating variables such as clarity and completeness. We expect (hope) the ratings on this scale will be higher following the completion of the inservice work on The Data Box.

C. Interview results

A two part interview was prepared to determine what the subset of teachers who volunteered for the total inservice felt assessment was all about and how they viewed the inservice. At Rhein Main and Hainerberg these interviews took place within a day or two following the workshop. At Karlsruhe and Heidelberg it was not possible to free up teachers during working hours for individual interviews. Therefore, in those two buildings, the volunteers met as a group after school was over. They were each given a copy of the questions, and filled in their own answers. Two members of our team were there to answer questions, probe, etc.

The data from this interview/questionnaire have been fully described in Wanous et. al. (1981). A general summary of the results is shown in Appendix E. There were no great surprises. In thinking about assessment most teachers think of testing and observation. Eighty six percent say they use assessment data in making instructional decisions. They felt a need for more standardized test data. They need it for new or transfer students. Barriers identified to the use of data included lack of access, availability, and relevancy. There was no apparent concern about quality of data.

The data from the interviews/questionnairs regarding the inservice is of limited value since only the workshop portion of the inservice had been completed. The post inservice interviews to be conducted in May should be much more valuable.

III. Post Implementation Evaluation Plans

Two members of our team will be returning to Germany in May to evaluate the total inservice activities. At this time we will readminister the cognitive assessment test and the workshop/Data Box evaluation scale. We will conduct interviews with all the participants in the total inservice. We will also gather the records (logs) the participants have been keeping as they have been working through the vignettes. These records will include reactions on such variables as completeness, clarity, relevance, and usefulness of The Data Box and its component parts.

We will be gathering similar data from London where a group of teachers have been working with The Data Box without a preceeding workshop.

The final evaluation, then, will be based on the total set of data discussed throughout this paper.

We will have data on the quality of the materials in The Data Box and

suggestions for improvement. We will also have some data comparing two delivery systems: one with a workshop and one without. Because of some difficulties in implementing our design we will be able to make only weak inferences about the relative merits of the delivery systems.

We hope to be able to tell you about our final conclusions next year!

References

Mehrens, William A. What Fundamental Concepts Shall We Stress? Handout at NCME Session 22: Integrating Assessment with Instruction - An NIE Study. Boston, April 10, 1980.

Porter, Andrew C. Evaluating Materials to Promote Better Use of Assessment Devices in Teacher Decision Making. Xerox, 1980.

Rudman, Herbert C., Mehrens, William A., Wanous, Donna S., Clark, Christopher M., and Porter, Andrew C. The Data Box Technical Manual, Michigan State University, 1980.

Van Susteren, Timothy J. Evaluating The Data Box. Unpublished paper, 1980.

Wanous, Donna S., Rudman, Herbert, Mehrens, William, Clark, Christopher, and Porter, Andrew. Integrating Assessment with Instruction: An Impossible Dream? Paper presented at the AERA Annual Meeting, Los Angeles, 1981.

Tables

1. Characteristics of the Tryout Population
2. Attitudes of the Tryout Population Towards the Components of The Data Box.
3. Reliability of the Instrument used in Evaluating The Data Box.
4. Pre Assessment Results.
5. A Comparison of Ratings of Workshops and The Data Box by Various Groups in North and South Germany, January, 1981.

TABLE 1

"What Were The Characteristics of the Tryout Population?"

Characteristic	Number	Percent
Teachers	34	92
Administrators	3	8
Population Center Type		
Urban	9	24
Suburban	11	30
Rural	17	46
Experience		
0-5 Years	17	46
6+ Years	20	54
Level of School		
Elementary	29	78
Middle or Junior High	8	22

TABLE 2

"What Were The Attitudes of the Tryout Population Towards The Components of The Data Box?"

Characteristics		Percent Agreement	Mean Rating	.95 Confidence Interval
1.0	THE VIGNETTES			
1.1	Problems Teachers Encounter	100	3.5	3.4-3.7
1.2	Problems Challenged Me Intellectually	95	3.4	3.2-3.6
1.3	Problems Presented Clearly	95	3.2	3.0-3.4
1.4	Problems Adequately Defined	97	3.2	3.0-3.4
1.5	Help To Develop Competence	94	3.2	3.0-3.4
1.6	Problems Are Practical	97	3.5	3.3-3.6
1.7	Problems Stimulate Discussion	95	3.4	3.2-3.5
1.8	Help Develop Useful Skills	97	3.3	3.1-3.5
1.9	Appropriate Amount of Material	68	2.8	2.6-3.0
1.10	Questions Focus on Problems	95	3.2	3.0-3.4
1.11	Questions Are Typical	84	3.1	2.8-3.3
1.12	Illustrative of Major Categories	100	3.3	3.1-3.4
1.13	Sufficient Information Provided	89	3.0	2.8-3.2
1.14	Opportunity For Self-Expression	94	3.2	3.0-3.3
1.15	Directions Carefully Explained	84	3.0	2.8-3.2
1.16	Situations I Face Often	97	3.5	3.3-3.7
2.0	THE DOCUMENT FILE			
2.1	Correspondence Is Realistic	95	3.1	3.0-3.3
2.2	Scores Are Familiar To Me	76	2.8	2.6-3.0
2.3	Scores Are Familiar to Most Teachers	43	2.4	2.2-2.6
2.4	Documents Are Sufficient To Analyze	81	2.8	2.6-3.0
2.5	Easy To Find Information	46	2.3	2.0-2.6
2.6	Report Forms Easily Understood	65	2.6	2.4-2.8
2.7	Documents Are Typical Of Those Used In Schools	70	2.8	2.6-3.1
2.8	No Type of Data Missing	65	2.6	2.4-2.8

TABLE 2 (Continued)

"What Were The Attitudes of the Try out Population Towards The Components of The Data Box?"

Characteristics		Percent Agreement	Mean Rating	.95 Confidence Interval
3.0	THE INDIVIDUAL PUPIL FOLDER			
3.1	Typical Permanent File	78	3.0	2.8-3.2
3.2	Folder Provides Adequate Information	86	3.0	2.8-3.1
3.3	More Information Than Found	57	2.7	2.4-2.9
3.4	Information Is Realistic	95	3.0	2.9-3.2
4.0	THE INTERPRETIVE MANUAL			
4.1	Manual Is Well Organized	95	3.3	3.1-3.5
4.2	Manual Provides Valuable Background Information	95	3.5	3.3-3.7
4.3	Manual Provides A Clear Explanation Of How To Use The Data Box	92	3.1	3.0-3.3
4.4	Style Easy To Understand	87	3.3	3.1-3.6
4.5	Faculty Could Conduct Meetings	89	3.1	2.9-3.3
4.6	Useful To Teachers, Counselors And Administrators	86	3.2	3.0-3.4
4.7	Clear Definition Of Terms	89	3.3	3.1-3.5
4.8	Use Manual As A Reference Tool	92	3.2	3.0-3.3
4.9	Understand Vignettes To Documents And To Measurement Concepts	73	2.8	2.6-3.0
4.10	Adequately Explains Goals of The Data Box	97	3.3	3.1-3.5
4.11	Help Plan In-Service Program	89	3.1	2.9-3.3
4.12	Problems Are Realistic	92	3.2	3.0-3.4
4.13	Length Is Adequate	87	3.1	2.9-3.3
4.14	Chapters Are Logical and Useful	92	3.2	3.0-3.4
5.0	SUMMING UP			
5.1	The Box Is Useful For Teachers	97	3.4	3.2-3.6
5.2	Help Teachers Supplement Their Intuition	97	3.4	3.2-3.5
5.3	The Box Should Be Used For In-Service Training	100	3.5	3.4-3.7

TABLE 2 (Continued)

"What Were The Attitudes of the Try out Population Towards The Components of The Data Box?"

	Characteristics	Percent Agreement	Mean Rating	.95 Confidence Interval
5.4	The Box Useful In Preservice Preparation	92	3.4	3.2-3.6
5.5	Approach Reflects Reality	87	3.2	3.0-3.4
5.6	Implement At System Level	46	2.6	2.4-2.8
5.7	Implement At Building Level	87	3.2	3.0-3.4
5.8	The Box In Classroom as Reference	84	3.2	2.9-3.5
5.9	The Box As Training for Administrators	95	3.4	3.2-3.6
5.10	The Box As Training for Counselors	97	3.4	3.2-3.6
5.11	Contexts Chosen Most Important	78	2.9	2.7-3.1

TABLE 3

"How Reliable Was The Instrument
Used In Evaluating The Data Box?"

Scale		Alpha
1.0	THE VIGNETTES	.85
2.0	THE DOCUMENT FILE	.51
3.0	THE INDIVIDUAL PUPIL FOLDER	.16
4.0	THE INTERPRETIVE MANUAL	.89
5.0	SUMMING UP	.80
	TOTAL	.92

Table 4

Pre Assessment Results (Max Score = 20)

Location	N	$\bar{X}$	Sx	High-Low Range	KR-20*	Se
Heidelberg	12	10.17	2.16	8 → 14	.19	1.94
Karlsruhe	42	10.93	2.69	5 → 17	.49	1.92
Rhein Main	11	12.00	2.04	9 → 15	.14	1.89
Hainerberg	8	11.75	2.96	7 → 16	.64	1.76
Total	73	11.05	2.57	5 → 17	.42	1.95

*Note: Given that this was a pretest and that the means are quite low (chance = 4.6) one would expect the realiability to be low.

A COMPARISON OF RATINGS OF WORKSHOPS AND THE DATA BOX
BY VARIOUS GROUPS IN NORTH AND SOUTH GERMANY, JANUARY, 1981

VARIABLES	GROUPS																	
	N. Germany Administrators			Total Teachers			Rhein-Main			Hainerberg			Karlsruhe			Heidelberg		
	$\bar{X}$	σ	N	$\bar{X}$	σ	N	$\bar{X}$	σ	N	$\bar{X}$	σ	N	$\bar{X}$	σ	N	$\bar{X}$	σ	N
1. Purpose	4.6*	.52	10	3.9	1.05	161	4.0	1.08	64	3.6	1.08	36	4.1	.96	40	4.1	1.07	21
2. Content	4.1	.74	10	4.0	.91	160	4.0	.87	64	3.9	1.00	35	4.0	.95	40	4.4	.74	21
3. Importance	4.6	.70	10	3.9	1.11	161	3.6	1.17	64	3.7	1.01	36	4.2	1.17	40	4.4	.67	21
4. Use	3.7	.82	10	3.4	1.18	160	3.3	1.18	64	3.4	1.08	36	3.4	1.20	39	4.0	1.24	21
5. Plans for Trial	3.7	.95	10	3.3	1.20	160	2.8	1.19	63	3.4	1.20	36	3.5	1.11	40	4.0	.95	21
6. Measurement Concepts	3.4	.79	7	3.1	1.12	159	3.0	1.12	62	2.9	1.07	36	3.1	1.07	40	3.9	.94	21
7. Attractive	3.8	1.0	10	3.3	.91	159	3.4	.96	63	3.1	.92	36	3.2	.84	40	3.8	.87	21
8. Clarity	4.1	.74	10	3.1	.93	159	3.1	1.00	63	3.1	.84	36	2.9	.82	40	3.4	1.08	21
9. Real	4.0	.67	10	3.2	1.10	158	3.1	1.23	63	3.1	.93	36	2.9	1.02	39	3.7	1.06	21
10. Organization	4.2	.63	10	3.4	1.01	159	3.6	1.00	63	3.3	.89	36	3.1	1.01	40	3.7	1.15	21
11. Complete	3.5	.71	10	3.1	.96	156	2.9	1.00	63	3.1	.99	36	2.9	.90	38	3.7	.81	20
12. Useful	3.6	.84	10	2.9	1.13	149	3.0	1.17	59	2.8	1.03	35	2.8	1.19	36	3.0	1.16	21

* 5 = Highest, 1 = Lowest

Appendices

A. Directions for Analysis

B. Questionnaire on Evaluating The Data Box

C. Pre Assessment Knowledge Test

D. Workshop Evaluation Form

E. Interview Questions and Results

F. Major Measurement Concepts

Appendix A

MICHIGAN STATE UNIVERSITY

COLLEGE OF EDUCATION
DEPARTMENT OF ADMINISTRATION AND HIGHER EDUCATION
ERICKSON HALL

EAST LANSING • MICHIGAN • 48824

DIRECTIONS FOR ANALYSIS

Thank you very much for your willingness to help us as we develop The Data Box in its final format. Your contributions to this project will mean a great deal to us and to the development of more effective approaches to the in-service preparation of teachers.

The following guidelines for the analysis of the materials contained in the packages given you should prove helpful to you as work your way through them:

(1) Turn to the Interpretive Manual and read Chapters 1, 2, 3, and 5. Chapter 4 contains the "solutions" to the problems posed in the vignettes you will read. But a careful reading of Chapters 1, 2, and 3 will give you a complete overview of The Data Box and will help you understand better what the various components are that you will be analyzing.

(2) Now quickly skim through the other packages given to you. You will find a set of vignette cards which present a situation and ask you to react to that situation in differents ways. You will also find a number of various documents that represent formal and informal assessment information. You will be using these throughout your analysis of the vignettes. Make yourself acquainted with them. You will also find an individual folder for a student named "Charlie Brown." This will also be useful to you in understanding those vignettes that may deal with him.

(3) Now turn to each vignette, read it carefully, think through solutions to the problems posed and then turn to Chapter 4 in the Interpretive Manual for our solutions. How adequate was the material presented to you? Were we practical or realistic? If we were not, in what ways were we not practical? Complete working your way through the vignettes and the various sections in Chapter 4 that you will find related to each vignette. This exercise will thus give you an opportunity to search through the various documents in the remainder of the materials.

(4) After you have read the manual, and worked your way through the vignettes and documents, turn to the questionnaire which accompanies these materials. As you begin to work your way through all of the materials, note the time that you started and when you have completed the questionnaire, note the time you have ended. Enter the total time spent in analyzing the materials and completing the questionnaire as a total number of hours and minutes. You should complete this analysis somewhere within one hour or two at the most. Your responses

to the second part of the questionnaire are as valuable (if not not even more so) than your responses to the first part. Please answer each of the items in the questionnaire.

(5) When you have completed all of your responses, please sign the front cover of the evaluation instrument so that we can reimburse you for the time spent.

One caution; please remember that this material is experimental and incomplete. There are two vignettes 1.1 and 4.1 which are complete but which do not have complete responses to them. They are being worked on now. All other vignettes are pretty much in place. We are also aware that the data in the documents need to be edited and revised somewhat. Our major interest at the moment is whether the system appears to be promising to you as a device for the in-service training of teachers, whether our vignettes and our interpretation of these vignettes are practical, and whether the system reflects the reality of your classroom. Your responses to this questionnaire will help give us some answers. We anticipate that when the system is complete it will contain a minimum of twenty vignettes and twenty interpretations to be found in Chapter 4. This our "toe in the water." Shall we take the full swim, or head for the bathhouse?

Good luck and thank you for this effort.

Appendix B

EVALUATING THE DATA BOX

Herbert C. Rudman, Timothy Van Susteren, William A. Mehrens, Donna S. Wanous,
Christopher M. Clark and Andrew C. Porter

Michigan State University

DIRECTIONS: Please place an "X" in the boxes after the characteristics that best apply to you:

001 Teacher
002 Administrator
003 Counselor

011 Urban
012 Suburban
013 Rural

021 0-5 Years of Teaching Experience
022 6 or More Years of Teaching Experience

031 Elementary School
032 Middle or Junior High School

This instrument is divided into two parts; a rating scale and an open-ended evaluation that permits you to deal with questions, issues and attributes not covered in the rating scale. Each part is categorized to reflect the major components of The Data Box. While the box itself is not available to you at the present time, the materials you will examine are draft manuscript copies of the materials that may be included in the final package. Please keep in mind that they are experimental drafts to be modified as needed based in part on your reactions and the reactions of other reviewers.

A sketch of the total package is attached to give some indication of how the final product will appear.

When you have completed your examination of the materials and have recorded your reactions to them please indicate in the space below the time spent by you in this activity.

We sincerely thank you for your help. It is through collaboration with teachers, administrators and counselors that the most effective kinds of help can be rendered. Your careful consideration is invaluable to this project.

Total Time Spent_____

Characteristics	Strongly Agree	Agree	Disagree	Strongly Disagree
	Place an "X" on the appropriate number			
1.0 THE VIGNETTES				
1.1 The vignettes address problems that teachers commonly encounter.	4	3	2	1
1.2 The problems presented in the vignettes challenged me intellectually.	4	3	2	1
1.3 The problems in the vignettes are presented clearly.	4	3	2	1
1.4 The situation or problem presented in the vignettes are adequately defined.	4	3	2	1
1.5 The vignettes would help a teacher to develop competence in dealing with a classroom-related problem.	4	3	2	1
1.6 The problems and situations posed in the vignettes are practical and reflect reality.	4	3	2	1
1.7 The problems are interesting and could stimulate discussions in a faculty meeting.	4	3	2	1
1.8 The vignettes could help teachers develop useful skills and techniques to deal with a variety of measurement problems.	4	3	2	1
1.9 Each vignette covers an appropriate amount of material.	4	3	2	1
1.10 The questions posed in the vignettes help the reader to focus on the problem.	4	3	2	1
1.11 The questions posed in the vignettes are typical of the questions teachers would pose themselves.	4	3	2	1
1.12 The vignettes are illustrative of the major categories in which they were placed.	4	3	2	1
1.13 Sufficient information is provided in the vignettes to allow for meaningful analysis of the problem presented.	4	3	2	1

Characteristics	Strongly Agree	Agree	Disagree	Strongly Disagree
	Place an "X" on the appropriate number			
1.0 THE VIGNETTES (continued)				
1.14 The vignettes provide an opportunity for self-expression.	4	3	2	1
1.15 Procedures and directions for proceeding are carefully explained in the vignettes.	4	3	2	1
1.16 I can recognize in the vignettes some situations that I often face in my classroom teaching.	4	3	2	1
2.0 THE DOCUMENT FILE				
2.1 The correspondence and memoranda are realistic and helpful in defining the problem.	4	3	2	1
2.2 Most of the scores and comments used on the profiles are familiar to me.	4	3	2	1
2.3 Most of the scores and comments used on the profiles are familiar to most teachers I know.	4	3	2	1
2.4 The documents in the file are sufficient information to analyze the vignette problems.	4	3	2	1
2.5 It was easy to find the information that I needed to analyze a particular vignette.	4	3	2	1
2.6 The data presented in the standardized test report forms was easily understood by me.	4	3	2	1
2.7 The documents found in the file are typical of those we use in the schools.	4	3	2	1
2.8 There does not appear to be any major type of data missing from this file that would represent the information I work with in my teaching.	4	3	2	1

Characteristics	Strongly Agree	Agree	Disagree	Strongly Disagree
	Place an "X" on the appropriate number			
3.0 THE INDIVIDUAL PUPIL FOLDER				
3.1 The folder is a typical representation of a student's permanent file.	4	3	2	1
3.2 The folder provides adequate information to analyze the problems posed in the vignettes.	4	3	2	1
3.3 Information stored in an individual's file generally contains more information than that found in this file.	4	3	2	1
3.4 The information contained in this file is realistic.	4	3	2	1
4.0 THE INTERPRETIVE MANUAL				
4.1 The manual is well organized.	4	3	2	1
4.2 The introduction to the manual provides the user with valuable background information.	4	3	2	1
4.3 The manual provides a clear explanation of how to use The Data Box.	4	3	2	1
4.4 The manual is written in a style that is easy to follow and understand.	4	3	2	1
4.5 The content of the manual is so written that a faculty could use it as a guide to conduct staff meetings built around test interpretation.	4	3	2	1
4.6 The manual provides a convincing explanation of the usefulness of The Box to teachers, counselors, and administrators.	4	3	2	1
4.7 The manual provided me a clear and precise definition of terms.	4	3	2	1

Characteristics	Strongly Agree	Agree	Disagree	Strongly Disagree
	Place an "X" on the appropriate number			
4.0 THE INTERPRETIVE MANUAL (continued)				
4.8 The manual is so written that I could use it as a reference tool for measurement problems I encounter in my teaching.	4	3	2	1
4.9 It is easy to understand the process described in the manual that keys vignettes to documents and to specific measurement concepts.	4	3	2	1
4.10 The manual does an adequate job of explaining the goals and objectives of The Data Box.	4	3	2	1
4.11 The interpretive manual could help teachers and administrators plan a continuing in-service program on the use and interpretation of assessment information.	4	3	2	1
4.12 The responses to the vignette problems provided in the manual are realistic and practical.	4	3	2	1
4.13 The length of the manual is adequate for the purposes it serves.	4	3	2	1
4.14 The chapters are presented in a logical and useful order.	4	3	2	1

.....

Now that you have completed your reactions to this first part of the evaluation, please share with us your responses to the following statements that would apply generally to the concept of The Data Box.

5.0 SUMMING UP				
5.1 The Data Box promises to be a useful tool for teachers.	4	3	2	1
5.2 This training package should help teachers supplement their intuition with information.	4	3	2	1
5.3 The Data Box should be used for in-service training of teachers.	4	3	2	1

Characteristics	Strongly Agree	Agree	Disagree	Strongly Disagree	
	Place an "X" on the appropriate number				
5.0 SUMMING UP (continued)					
5.4 The Data Box has a useful place in the preservice preparation of teachers.	4	3	2	1	
5.5 This approach to helping teachers is practical and reflects the reality I confront during the year.	4	3	2	1	
5.6 This approach can best be implemented at the system level.	4	3	2	1	
5.7 This approach can best be implemented at the building level.	4	3	2	1	
5.8 I would find it useful to have The Box in my classroom for use as a reference when needed.	4	3	2	1	
5.9 The Data Box would serve well as a training device for administrators.	4	3	2	1	
5.10 The Data Box would serve well as a training device for counselors.	4	3	2	1	
5.11 The eight contexts within which the vignettes are categorized represent the eight most important teaching concepts.	4	3	2	1	

.....

This second part of the evaluation is open-ended so that you may respond to your recent introduction to The Data Box in ways we could not anticipate. We have organized this section to reflect the five components covered in the first part. Please feel free to respond in any way that you think appropriate.

Have we asked the important questions? Are there situations, documents, interpretations, and the like that we have not included that you think should be included? Do you agree with our interpretations of the vignettes in Chapter IV? If you differ please show us how you would interpret the "solutions" (your way may be at least as useful--if not more so--than ours).

1.0 THE VIGNETTES

2.0 THE DOCUMENT FILE

3.0 THE INDIVIDUAL PUPIL FOLDER

4.0 THE INTERPRETIVE MANUAL

5.0 SUMMING UP

Appendix C

Inservice PreAssessment

On the following pages are twenty questions covering some measurement concepts. Read each item and select the best answer. Fill in the space on the answer sheet provided.

Concept 4.0

1. In order to study behavior adequately
   A. there must be sufficient representative observations.
   B. interpretation must be separate from observation.
   C. the subject must not know he is being observed.
   D. the observer must know the subject quite well.
   E. All of the above.

2. In keeping anecdotal records, the teacher should be primarily concerned about recording
   A. information that can be corroborated by other observers.
   B. behaviors that may be indicative of maladjustment.
   C. significant incidents in the life of the child.
   D. actions that are disruptive of the class.
   E. interpretations of the child's behaviors.

3. The most easily defensible use of standardized tests in modern schools is for
   A. accreditation of schools.
   B. maintenance of standards.
   C. motivation of study.
   D. teacher evaluation.
   E. pupil guidance.

4. Mr. McVey needs a test to measure instructional objectives unique to him. He should probably use a(n)
   A. essay test.
   B. objective test.
   C. teacher-made test.
   D. standardized test.

5. In terms of content sampling, standardized achievement tests are usually designed to assess achievement over
   A. the student's complete cumulative education.
   B. one or two units of work.
   C. one or more year's work.
   D. one semester's work.

Concept 5.0

6. The standard error of measurement is useful for
   A. reporting an individual's score within a band of the score range.
   B. converting individual raw scores to percentile ranks.
   C. reporting a group's average score.
   D. comparing group differences.

7. Mr. Smith compares the items on the Cooperative Mathematics Test with his course objectives. With what kind of test validity is he concerned?
   A. Concurrent
   B. Construct
   C. Content
   D. Predictive

Concept 6.0

8. Which type of item format provides greatest content sampling in the shortest possible testing time?
   A. Essay
   B. Matching
   C. Multiple-choice
   D. Short-answer
   E. True-false

9. Attention to the administration and scoring of a test is important because
   A. effort expended to construct a good test may be wasted if extraneous factors prevent results from reflecting true achievement.
   B. failure to do so will decrease the variance of the test and cause the group to appear more homogenous than it really is.
   C. a larger standard error will be obtained if instructions are not clear and this will result in lower reliability.
   D. the validity of the test will be lowered if students make unnecessary errors or the scoring process is not accurate.

10. In evaluating an achievement test for possible use in your school, the most important issue is whether the test
   A. is based upon a thorough survey of teaching practices over the whole country.
   B. matches in content the objectives of instruction in the school.
   C. has sufficiently high predictive validity.
   D. is convenient to administer and score.
   E. has sufficiently high reliability.

11. In attempting to decide whether a specific achievement test is appropriate for the measurement of the goals for which he has taught, the teacher should be guided primarily by
    A. test reviews in professional journals.
    B. Buros' Mental Measurements Yearbook.
    C. the content of the task.
    D. the Education Index.
    E. the test manual.

12. The first step in selecting a standardized test for use in school is to
    A. consult other schools concerning the types of tests they have found useful.
    B. consult Buros' Mental Measurements Yearbook.
    C. determine the type of information needed.
    D. review test publishers' catalogues.

Concept 7.0

13. What is the point on the scale that has an equal number of scores above and below it?
    A. Mean
    B. Median
    C. Mode
    D. Average deviation
    E. Standard deviation

14. To determine how useful a test can be in decision making, an administrator should consider
    A. the availability of the other data.
    B. the cost of faulty decisions.
    C. the selection ratio.
    D. the success ratio.
    E. All of the above.

15. A fourth-grade teacher found that half the class scored below average on a reading test given at the end of the year. Assuming the teacher was teaching reading all year, what should be concluded?
    A. No conclusion should be made until beginning readiness test data are correlated with final achievement test data to determine validity.
    B. A readiness test should have been used at the beginning of the year in addition to a general aptitude test.
    C. Gain scores between the readiness and the achievement test should first be determined.
    D. These results could be normally expected.

16. Ms. Ensign wishes to evaluate her new reading program. What role should standardized test results play?
    A. They should be used only if supplemented by teacher-made tests designed to measure Ms. Ensign's local objectives.
    B. They should be used only if the data are supplemented by the pupil's opinions.
    C. They could be used without any additional data.
    D. They should not be used in the evaluation.

17. Which one of the following exemplifies norm-referenced measurement?
    A. Alice got 63 out of 100 points on the anthropology final.
    B. Joe was admitted to honors college this year.
    C. Ed earned a 4.0 in his measurement class.
    D. Susan's score was average for her class.

Questions 18 and 19 are based on this partial printout of achievement test results for one senior high school student, Matthew. NOTE: RS = raw score, PR = percentile rank, and T = T score.

Arithmetic Scores					Reading Scores				
Matthew's	Westview		U.S.		Matthew's	Westview		U.S.	
RS	PR	T	PR	T	RS	PR	T	PR	T
36	7	35	58	52	22	79	58	99	60

18. How does Matthew compare to the U.S. average in arithmetic achievement?
    A. Extremely far above the U.S. average
    B. Somewhat above the national average
    C. Right at the national average
    D. Somewhat below the national average

19. How do children in Westview seem to compare with children nationally?
    A. Westview children do somewhat better than U.S. children.
    B. Westview pupils are better in reading but worse in math.
    C. Westview pupils are average in reading but poorer in arithmetic.
    D. Overall, Westview children do worse than U.S. average.

Concept 8.0

20. Feedback about test performance is most useful when it is
    A. immediate and specific.
    B. immediate and general.
    C. delayed and specific.
    D. delayed and general.

Appendix D

Michigan State University
College of Education
Workshop on
Integrating Assessment With Instruction

Workshop Evaluation Form

I. Workshop Objectives. Indicate the extent to which the workshop met the following stated objectives for you:

	Fully Met 5	4	3	2	Not At All 1
1. Introduced me to the purposes of The Data Box.	___	___	___	___	___
2. Familiarized me with the contents of The Data Box.	___	___	___	___	___
3. Reminded me of the importance of using data in instructional decision-making.	___	___	___	___	___
4. Demonstrated how The Data Box can be used to integrate assessment information with instructional decisions.	___	___	___	___	___
5. Clearly informed me of plans for the field trial of The Data Box.	___	___	___	___	___
6. Taught me how particular measurement concepts can be used in interpreting assessment data.	___	___	___	___	___

II. Data Box Materials. Rate The Data Box materials that you used in the workshop in comparison with other teacher aids or training materials with which you have worked.

	Superior 5	Above Average 4	Average 3	Below Average 2	Unsatisfactory 1
1. Attractiveness	___	___	___	___	___
2. Clarity	___	___	___	___	___
3. Realism	___	___	___	___	___
4. Organization	___	___	___	___	___
5. Completeness	___	___	___	___	___
6. Potential Usefulness for teachers	___	___	___	___	___

III. Comments. In the space below, please write any comments you have that explain or elaborate your ratings on the above items or that provide additional feedback that would be helpful to our staff. Continue your comments on the reverse side of this form if necessary.

Appendix E

INTERVIEW QUESTIONS AND RESULTS

I. ASSESSMENT

1. When you think about data or kinds of assessment data you use to make instructional decisions in your classroom, how would you describe that data?

Teachers generally think of testing (standardized or teacher-made) and observation.

(1a.) Can you give me some specific examples?

No real surprises except for some of the physical characteristics they included.

2. Would you say that most of your instructional decisions are made using some type of assessment data?

86% said yes

(2a.) Why or why not?

Usually for grouping and placement - especially in reading and math. Didn't use if the folders were late or they taught early elementary grades.

3. In what areas of instructional decisions do you think you might use assessment data more frequently?

They would use it to further assess math and reading abilities along with grouping, planning, and grading. There was a heavy emphasis on need for more standardized test data.

(3a.) For example, would it help if you had easier access to the data, it was better organized, or you had more knowledge about what the data meant?

Many teachers stated that the data were not accessible, they were not processed on time and needed to be better organized. There was a sense that much of the information was "specialist" controlled.

4. Can you identify any barriers to the use of assessment data in your classroom decision making?

The barriers were the time that it would take to coordinate the data, lack of access, availability of certain data and the relevancy of the data (e.g., DODDS own tests).

5. In the affective area, what kinds of assessment data do you use?

Teachers used other people's opinion or data most (38%). This includes specialists, other teachers, parents and other children. There was virtually no formal assessment done although the teachers clearly thought this was "the important" area. Nineteen percent of the teachers didn't use affective data or couldn't identify it.

6. Under what conditions or in what circumstances, do you find the results from standardized tests useful?

Importance was placed on test data at the beginning of the year or for a transfer student. They used it to determine the overall pattern of a student's achievement, for placement and grading.

7. In your classroom, are there assessment problems for which you could use some help?

28% of teachers couldn't identify a problem. The two problems most frequently cited were diagnosis of children with learning disabilities and detailed assessment of skills/problems of slow readers.

II. THE INSERVICE

1. Did you find the vignettes reflected realistic classroom problems?

yes - 85% yes & no - 12% no - 3%

(1a.) Why or why not?

#1 Could identify with problem 48%
#2 Really helped to have affective component 7%

2. If you could add to the vignettes, what critical incidents would you include?

1. drug related problem
2. DODDS specific problem - military specific
3. classroom management
4. reading placement
5. how to deal with students when they return from specialist
6. heterogeneous groups
7. parental problems
8. children with mixed profile of skills
9. lower elementary grades
10. health problems

3. Were the documents in the file similar to those you have access to for your class?

yes - 44% not really - 25% some - 28% no - 3%

4. Did you find the solutions to the vignettes acceptable?

yes - 88% no - 12%

(4a.) Do you agree (match) with them most of the time?

yes - 66% not quite - 28% no - 6%

5. How would you evaluate the format and content of the materials?

positive comments - 48%
negative comments and suggestions - 52%

PROBES: too simplistic, unworkable, coherent unit...

Negative comments

1. Needed more time
2. Too much material
3. Needed more individual data - including reading scores
4. More explanation about how to read forms
5. Include more grade levels
6. More teacher observation data
7. Hard to read

Positive Comments

1. Well organized
2. Logical
3. Realistic
4. Taught me about different kinds of data
5. Good format
6. Easy to understand

6. Did you like the overall staff development plan for these materials?

yes - 42% not sure - 3% negative - 55%

Time and expectations = most common negative comment.

Suggestions included less time on introduction and do more than just one vignette.

7. Do you believe that you have had an adequate amount of instruction to understand the necessary measurement concepts?

yes - 55% no - 42% ? - 3% Liked group interaction.

(8a) How?

Possible Growth Areas

1. Become more reflective
2. Enhance my teaching
3. Become more systematic
4. Would like to develop own self assessment of my teaching
5. Teacher sharing with common knowledge
6. Enhance my ongoing learning

8. Would you change your own teaching given this experience?

yes - 39% not sure - 33% other suggestions - 28%

Major Measurement Concepts Teachers Need to Know

I. Recognize that educational decisions should be based on data.

II. Understand that teachers need to know what data to gather, how (where) to gather it, how to use it in their own decision making, and how to disseminate the information to other decision makers.

III. Know about various sources of data such as the following:

1. parents
2. previous teachers
3. support staff
4. physicians
5. student
6. current teacher himself/herself

IV. Know about data collection processes including the following:

1. previous records (e.g. CA. 60)
2. interview with above sources
3. observations (formal & informal)
4. testing (teacher & standardized)

V. Understand data characteristics

1. Relevance (a. content validity, b. predictive validity)
2. Consistency
   a) types (causes) of random errors
   b) standard error of measurement and band interpretation
3. Costs: Financial & Psychological
   includes ease of gathering, invasion of privacy, etc.

VI. Interaction of Data Characteristics with Source & Process

1. Know how to build instruments with appropriate data characteristics.
2. Know how to choose instruments with appropriate data characteristics.
3. Know how to analyze instruments with respect to appropriate data characteristics.

VII. Know how to use data in Decision Making

1. Know how to combine (weight) data--Statistical advantage over clinical
2. Know how to convert data to common scale
3. Understand how to interpret various types of scores such as percentiles, T scores, stanines, grade equivalents, and deviation IQ scores.
4. Understand concepts of two types of errors and their relative costs.
5. Recognize that there is no reason to diagnose unless it leads to differential treatment.

VIII. Know how to disseminate information

1. to children
2. to parents
3. to other educators
4. to the general public

IX. Understand what information should be disseminated to the above named groups.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Issues in Inservice Introduction of the Data Box[1,2]

Christohper M. Clark, William Mehrens, Andrew C. Porter,

Herbert C. Rudman, and Donna Wanous

Michigan State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Rudman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 343 (4 of 4)

---

[1]Paper presented to the National Council on Measurement in Education, Los Angeles, California, April 15, 1981.

[2]This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions in this publication do not necessarily reflect the position, policy or endorsement of the National Institute of Education.

Inservice Introduction of the Data Box

After sixteen months of research and development the moment of truth had arrived. At last, in January, 1981, we were ready to offer the Data Box to a group of teachers for full implementation and field evaluation. Our plan was simple and straightforward: (1) present an overview of the Data Box to regional and building administrators to gain their endorsement and support, (2) spend two days at each school site to learn about the local situation and to establish credibility with teachers, (3) present a one-day workshop on the content and methods of using the Data Box, and (4) to enlist volunteer teachers who were willing to work through the entire set of thirteen vignettes over the succeeding three-month period.

This plan was a good one, and we were able to execute it without major modifications in four different schools in the DODDS System in West Germany. But what made this inservice implementation phase interesting and challenging for us were the issues and circumstances that we could not have possibly predicted--issues and circumstances that made each school situation unique and that required continuous adjustment and fine tuning of our basic plan. This paper is about those issues and the ways in which we responded to them.

I have grouped the issues that we confronted in implementation of the Data Box inservice workshops into two broad categories: Teacher Morale and Teacher Needs. As we examine each of these clusters, it is important to be aware that we were asking teachers to do something rather significant and demanding as they worked through the Data Box: to change the way that they think about their use of data in decision making. Furthermore, we asked this particular

group of teachers to keep track of changes in their thinking and to compile ideas for revision and improvement of the materials. Thus, volunteering to participate in this project represented a substantial, demanding, and potentially threatening commitment.

Teacher Morale

Teacher morale had an important influence on the way that the Data Box was received because participation in the project entailed extra teacher time and effort. My hypothesis is that, in situations in which teacher morale is high and teachers are feeling professionally competent and in control of the teaching situation, they are willing to take on a challenging professional development task. When morale is low, teachers are less willing to enter into such situations.

Among the DODDS teachers to whom we introduced the Data Box, morale varied from high to low. We have no specific measure of teacher morale to draw upon here, but during our first week in Germany we learned of a number of conditions that were having an impact on teacher morale. First, as it the case with many school systems today, DODDS is in the throes of a severe budget cutting process. Shortly before we arrived, the teachers learned that their budget for books and supplies for the next school year would be cut in half. Ours was the first delegation from CONUS to visit DODDS Germany schools after this announcement was made, and our group included a DODDS staff member from the Washington, D.C. headquarters. Needless to say, the teachers and principals had this severe budget cut on their minds and were tempted to use the occasion of our visit as an opportunity to "tell off" the Washington bureaucrats.

A second factor affecting teacher morale was that, in their judgment, the quality of their students had been declining steadily for the past several years. Several teachers associated this apparent decline with the "all-volunteer army" concept and the attendant drop in formal educational level of the average recruit. As we have seen in other public school systems in which the school clientele is changing (by adding lower SES students), there was a tendency to romanticize the "good old da' especially among teachers with long tenure in the system. There also seemed to be an increased focus on classroom management as a response to student behavior problems not encountered in the past. The important point here is that teachers perceived their jobs as becoming more demanding, less rewarding, and less appreciated by the parents of their students and by the administrators higher up in the system.

A third morale factor was that mobility within the DODDS System had declined during the past few years. No longer could a teacher request a direct transfer from, say, Germany to England and expect this wish to be granted. Conventional wisdom now held that a teacher would have to spend two years or so at a remote post in the system (e.g., Iceland) while enroute to a more desirable station such as England. The net result was that some of the teachers who spoke with us felt trapped in their present assignment, with little hope of relief.

It is an understatement to say that this set of conditions is not conducive to high teacher morale. Yet I do not want to paint too bleak a picture. Some teachers were clearly happy with their assignments, confident in their ability to "do more with less" as the budget shrank, and optimistic about being able to teach each individual

effectively. Yet I do think that the tenor of the times was more discouraging than encouraging when we arrived in Germany and asked DODDS teachers to step back from their day-to-day problems and think about the ways that they use assessment information in their instructional decision making.

Teacher Needs

Every salesperson or entrepreneur knows that one formula for success is "find a need and fill it." Our development of the Data Box was a response to a jointly felt need on the part of National Institute of Education staff and DODDS Washington central staff. The NIE staff was primarily interested in basic research and development on teachers' uses of assessment information in instruction. DODDS staff members shared this interest, and also had a more specific and applied need, namely, to train teachers in the DODDS System to make good use of the formal assessment information that was being made available through a newly launched criterion-referenced diagnostic and prescriptive testing program and through a new basic skills testing program.

We certainly believe that we met successfully the NIE/DODDS needs expressed in the request for proposals issued eighteen months ago. The literature review and the Data Box itself provide new knowledge about and new methods for teacher training in interpretation and use of assessment information. We think that we have more than satisfied the original need. But sometimes the needs of the sponsors of a research and development program are different from those of the eventual consumers or users of the research and development products.

What did we learn from the inservice introduction of the Data Box to teachers in the field that bears on this issue?

In a word, we learned that teacher needs were different from the needs of system administrators. We received our first hint of this discrepancy from a DODDS regional staff member during our second day in Germany. He predicted that we would encounter resistance in our workshops because the whole program was seen locally as an imposition from Washington. Apparently, all previous inservice training in these schools resulted from requests initiated by teachers themselves, and reflected specific locally-felt needs. Our visit, in contrast, was initiated, scheduled, and organized by DODDS Washington staff, and the topic on which we were offering training did not connect in an obvious way with priority needs expressed by teachers.

A related issue that affected the reception that we and the Data Box received was that the two new testing programs being implemented were, at this early stage of development, seen as nuisances by the teachers rather than as needed aids. The Rasch-model criterion referenced tests and their associated summary reports were complex and not yet completely synchronized with curricula. The Basic Skills tests had been administered in the Spring of the previous school year and therefore were useless for diagnostic purposes (especially given the high student transience rate in these schools). In fact, the Basic Skills Testing Program was seen by some teachers as a potentially threatening teacher accountability device rather than as a useful source of data for teachers' instructional decisions. In short, to the degree that the Data Box was associated with these somewhat unpopular testing programs, we had an uphill battle convincing teachers that this training met their own needs.

There was yet another sense in which the Data Box was not exactly the answer to a teacher's prayers. Implicit in the Data Box design is the idea that teachers should always be striving to improve their performance as professionals, especially with regard to systematic use and interpretation of data. This view is certainly shared by most teacher educators, researchers on teaching, school administrators, and leaders of professional organizations. But, to a great extent, I do not believe that this veiw was a high priority operating principle for the teachers we worked with. Certainly, they were interested in professional improvements if those improvements produced immediate solutions to practical problems with which they were struggling. But increased professional competence for its own sake, separated from immediate practical problems, was a luxury that few had the time or inclination to indulge.

With regard to data interpretation and use, our teachers already had solved their problem by using the data that were already at hand, together with heavy reliance on their own experience and judgment, to make the decisions that they needed to make. Data that were too difficult to interpret or to obtain were either ignored or referred to the school psychologist or to other specialists more expert in these matters. These teachers, like most reasonable people in demanding situations, were "satisficers" rather than "optimizers." That is, they were more concerned with getting the job done adequately with a reasonable amount of effort than with doing the best of all possible jobs.

Conclusions

I could go on with more commentary about teacher morale, teacher needs, and teacher expectations and how they influenced the reception of the Data Box. Suffice it to say that, in the face of the issues discussed above and others unreported here, we were able to enlist approximately sixty teacher volunteers to work through the full formative evaluation of the Data Box. Our personal persuasiveness helped. The unexpected incentive of twelve professional recertification credits helped. The invitation to become collaborative teacher-researchers with us helped. And, most of all, the leadership and cooperation of the school principals and a few especially talented and enthusiastic teachers helped us to "sell" the Data Box in Germany. Low teacher morale and an imperfect match with teacher needs and expectations did not sabotage the inservice workshops, but they certainly made them challenging for us.